Available online at www.sciencedirect.com

ScienceDirect

Procedia - Social and Behavioral Sciences 97 (2013) 238 – 245

Procedia
Social and Behavioral Sciences

The 9th International Conference on Cognitive Science

# Effect of interleaving exemplars presented as auditory text on long-term retention in inductive learning

Norehan Zulkiply*

*Faculty of Cognitive Sciences and Human Development, Universiti Malaysia Sarawak,Kota Samarahan, 94300 Sarawak,Malaysia*

**Abstract**

The present study investigated the exemplar interleaving effect on long-term retention in the context of inductive learning. In particular, an experiment using aurally presented texts was conducted to examine whether or not the benefits of interleaving of exemplars from several categories vary with retention interval in inductive learning. Participants studied a series of aurally presented texts from several categories which were presented massed and interleaved, and later their induction was tested either shortly after the study phase (short-term retention) or after a week's delay (long-term retention). Consistent with findings from previous studies, the interleaving effect was found in the short-term retention condition, and crucially, the present study extended the generalisability of the interleaving effect in the long-term retention condition to aurally presented texts. Interestingly, participants judged massing to be more effective than spacing (interleaving) though their actual performance showed the opposite. The present study provides further evidence that interleaved exemplars have considerable potential in improving inductive learning in the long term.



*Keywords*: interleaving effect; inductive learning; category learning; category induction; spacing effect

## 1. Introduction

It is an established experimental finding that memory retention for spaced repeated items is better than massed repeated items [1, 2, 3]. The finding of improved memory for spaced repetitions, called the spacing effect, has been documented in a broad range of memory tasks with many different types of study materials [1, 2, 4]. The spacing effect is not only found when the test is given shortly after the study phase, which measures short-term memory retention [5, 6], but also when the test is given after a delay interval which measures long-term memory retention ranging from days [7, 8] to months [9].

In the context of inductive learning, the spacing effect was found over the brief retention interval (which measures short-term retention) in a number of studies [11-17]. Inductive learning is the process of learning by examples. Induction is concerned with inferring knowledge from an incomplete set of observations, and this contrasts with deduction, where the learner formulates regularities observed in a complete set of data [10]. In past studies that found the spacing effect, the focus was on inductive learning that occurs during category learning which requires students to work from specific exemplars and derive general concepts or categories from those exemplars.

* Corresponding author. Tel.: +6-082-581538; fax: +6-082-581567.
*E-mail address:* znorehan@fcs.unimas.my


doi:10.1016/j.sbspro.2013.10.228

In a typical study that investigates the spacing effect in inductive learning, participants are asked to learn exemplars from several categories which are presented with a variable degree of spacing between exemplars, and at the end of the session induction is tested on a set of novel exemplars from the same categories learnt in the study phase.

Later reports by Kang and Pashler [11] and Zulkiply and Burt [12] revealed that the factor that critically contributes to the benefits of spacing for induction is the interleaving of exemplars from different categories and not the temporal spacing of exemplars from the same categories. The spacing manipulation used in Kang and Pashler [11] and Zulkiply and Burt [12] was similar to those used in previous studies [11,13,14,15,16,17], thus it is possible that the benefits of spaced presentation found in these studies was due to the interleaving effect. In addition, in previous studies that found the spacing (interleaving) effect [11,13,14,15,16,17], induction was tested shortly after the study phase, measuring the effect for short-term retention. It is not really known whether or not interleaved presentation of exemplars from the same categories affects long-term retention. In a working paper by Zulkiply and Burt [18], the initial evidence of the interleaving effect in inductive learning in the long-term was provided using paintings and visually presented texts. Besides pictures and visually presented texts, aurally presented texts are also commonly used to present information, as in a lecture or speech, and understanding which textual presentation method (massed or interleaved) enhances long-term retention will be benefial to students. It would be interesting to find out if the benefits of the interleaving effect in inductive learning could also be found over longer retention interval, if aurally presented texts are used as the learning materials.

The present study examined whether or not the benefits of interleaving of exemplars from several categories vary with retention interval in inductive learning when aurally presented texts were used as the learning materials. In specific, the present study aimed to find out if the benefits of interleaving in the long-term can be generalized to aurally presented texts.

## 2. Method

### *2.1. Participants and design*

The participants in this study were 40 undergraduate students (20 females, 20 males). The design of the experiment was a 2 (Presentation style: Massed vs. Interleaved) x 2 (Retention type: Short-term vs. Long-term) x 3 (Test block: Blocks 1- 3) mixed-factorial design. Retention type was varied between-participants, while presentation style and test block were varied within-participants. There were four steps involved in the experimental manipulation: presentation (study) phase, distractor task phase, test phase and question phase. In the presentation phase, 18 case studies from six psychopathological categories were arranged in six learning blocks (three blocks for massed presentation; three blocks for interleaved presentation). The order of the blocks was 'MIMIMI' (M for massed; I for interleaved). The assignment of psychological disorders to condition (massed vs. interleaved) was counterbalanced over two lists. Two versions of each list were produced in which there was a different assignment of disorders to blocks. Thus, there were four lists in total. In the test phase, the 18 new cases from the six psychopathological categories were arranged in three test blocks. Each block consisted of one new case from each category, presented in a fixed order across participants. In the test phase, the 18 new cases from the six psychopathological categories were arranged in three test blocks. Each block consisted of one new case from each category, presented in a fixed order across participants.

### *2.2. Materials*

The materials were 36 case studies developed from six categories of psychopathological disorders as in Zulkiply et al. (2012) (see Appendix A for samples of case studies). As noted, 18 cases were used in the presentation/study phase (three cases per category) and another 18 cases were used in the test phase (three cases per category). The psychopathological disorder categories used were identifed by nonsense names to minimise the effects of participants' prior assumptions and expectations. Table 1 below illustrates the six disorder categories chosen as the basis of the case studies as well as the novel names assigned to each of the categories. Each case study was between 100-120 words in length and incorporated a description of a few symptons representative of the four factors of symptons in general: Cognitive, Behavioural, Emotional, Physical. Each case study was recorded onto the computer

by an Australian speaker, with a maximum duration of 40 seconds. All the case studies used in the experiment were pilot tested by ten Clinical PhD students.

Table 1. Categories and novel names of psychopathological disorders used to develop the case studies

| Categories of psychopathological disorders | Novel names assigned |
|---|---|
| Obsessive Compulsive Disorder | Duv |
| Phobia Disorder | Baj |
| Schizophrenia Disorder | Tem |
| Attention Deficit Disorder (Inattentive Type) | Pliq |
| Attention Deficit Disorder (Hyperactive and Impulsive Type) | Hix |
| Depression Disorder | Wos |

### *2.3. Procedure*

Participants were randomly assigned to either the short-term retention condition or the long-term retention condition. Participants first were instructed about the nature of the experiment before they entered the presentation phase which was subject to experimental manipulation. In the presentation phase, participants were asked to listen to 18 case studies and study the nature of the cases. Each case was aurally presented to the participants for approximately 40 seconds. While participants listened to each of the case studies, the novel label of the corresponding category was displayed on the computer screen. Next, participants were asked to complete a distractor task, during which they were asked to count backward by 3s starting from 723, for 15 seconds, while typing the numbers in the designated box on the computer screen. For the test phase, participants in the short-term retention condition were tested immediately after the end of the distractor task, whereas participants in the long-term retention condition were tested a week later. During the test phase, participants were presented with 18 new cases which they had not listened to before, and they had to identify to which category each case belonged. Participants listened to one case study at a time, with seven buttons displayed on the computer screen. Six of the buttons were labelled with the names of the case categories and one button was labelled 'I don't know'. Participants selected the category for each case study by clicking their computer's mouse on one of the seven buttons. Feedback was given after each response. If participants clicked on a correct category name for each displayed case study, the word 'correct'would appear in the middle of the computer screen. If they were wrong, the correct category name would be presented on the computer screen. Participants completed the test phase at their own pace. After the test phase, participants read a description about the meanings of the terms 'massed' and 'spaced (i.e., interleaved)" on the computer screen. They were asked, 'Which option do you think helped you learn more?' and were provided with three possible answers: 'massed', 'about the same', or 'spaced'. This question phase ended the experimental manipulation. Participation in the experiment took approximately 40 minutes and participants were debriefed about the experiment they had just participated in before they left the experimental room.

## 3. Results

The data were analysed with a three-way repeated measure ANOVA. As shown in both Figure 1 and Figure 2, participants' performance in interleaved study was significantly better than their performance in massed study, $F$ (1, 38) = 3.21, $p$ = .034, $\eta_P^2$ = .21, and participants' accuracy also increased significantly across test blocks, $F$(1.7, 64.66)= 3.28, $p$ = .023, $\eta_P^2$ = .13. The main effect of retention interval was also significant, $F$(1, 38) = 15.343, p < .001, $\eta_P^2$ = .35 . In addition, performance after the short retention interval was better than performance after the long retention interval. All the two-way and three-way interaction effects were not significant (presentation style and retention type, $F$(1, 38) = .056, $p$ = .870; test block and retention type, $F$(1.7,64.66) = .31, $p$ = .830; presentation style and test block, $F$(2, 76) = .01, $p$ = .840; presentation style, test block and retention type, $F$(2, 76) = .33, $p$ = .831)

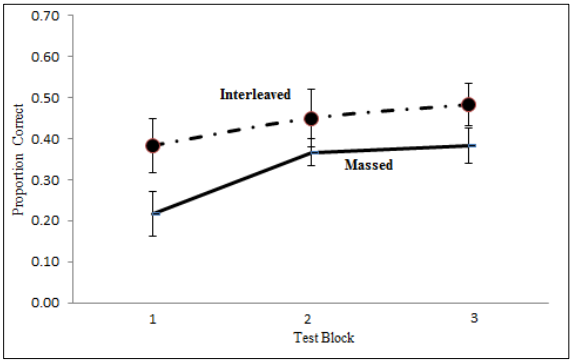


Fig. 1. Proportion of psychopathological categories selected correctly on the test in the short-term retention condition in the experiment as a function of presentation condition (interleaved or massed) and test block. Error bars represent standard errors.

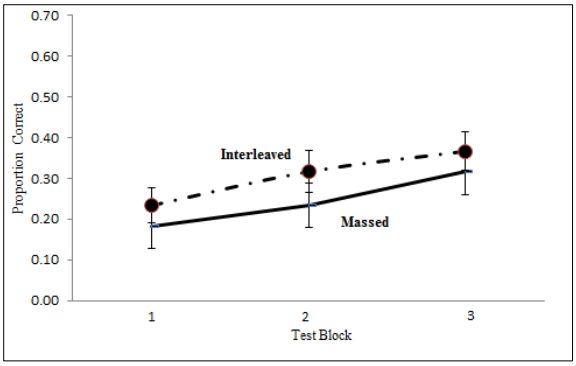


Fig. 2. Proportion of psychopathological categories selected correctly on the test in the long-term retention condition in the experiment as a function of presentation condition (interleaved or massed) and test block. Error bars represent standard errors.

With regards to participants' responses to the questionnaire administered after the test, that is, on which study presentation helped them learn more, a similar preference for massed presentation was observed in both retention conditions. In the short-term retention condition, a one-way Chi-square analysis showed a significant difference among the three judgement options, $\chi 2$ (2, $N=20$) = 12.9, $p$ = .007. Of a total of 20 participants, a majority of 14 (70%) claimed massed presentation was better, 5 (25%) claimed spaced and 1 (5%) judged that both massed and spaced (interleaving) contributed equally in helping them to learn during the learning phase, regardless of their performance in the two conditions . In terms of categorisation performance, 13 (65%) of the participants performed better in spaced (interleaved) condition, 5 (25%) performed better in massed condition and 2 (10%) performed equally in the two conditions.

Similarly, in the long-term retention condition, of a total of 20 participants, a majority of 14 (70%) participants claimed massed was more effective, 2 (10%) claimed spaced (interleaved) and another 4 (20%) judged the two conditions equally effective, regardless of their performance in the two conditions. The result of a one-way Chi-square analysis conducted on the judgement data confirmed our prediction, $\chi 2$ (2, $N=20$) = 14.1, $p$ =. 008. In terms of categorisation performance, 11 (55%) of the participants performed better in spaced (interleaved) condition, 3 (15%) performed better in massed condition and 6(30%) performed equally in the two conditions.

## 4. Discussion and Conclusion

The present finding that short-term retention benefits from the interleaved presentation in inductive learning was consistent with findings from past studies [11,13,14,15,16,17]. In addition, the finding that the interleaving of aurally presented texts from different categories also affected long-term retention was consistent with findings reported in Zulkiply and Burt's study [18], which used paintings and visually presented texts as the learning materials. Crucially, the present finding extended the generality of the interleaving effect to aurally presented texts over longer retention interval in the context of inductive learning --positioning different individual auditory exemplars apart in time aids in the learning of categories in the long term. Interleaving had similar effects on induction in the long- and short-retention groups, eventhough performance was generally better when induction was tested shortly after the study phase than after a one week's delay (and overall accuracy at the beginning of the test session was lower for the long-retention group.)

The first possible account for the interleaving effect (which resulted from the spaced presentation) in induction concerns the association between induction and discrimination processes [e.g., 13]. The interleaving of exemplars from different categories in the spaced condition might have fostered and enhanced discrimination learning, allowing participants to compare and contrast the auditory text exemplars from different categories and to notice the different natures of the case studies —this would be expected to assist participants in understanding the nature of each psychopathological category. Several authors have suggested that interleaving promotes the apprehension of points of contrasts among exemplars, making these differences among the categories more salient [e.g., 11,19]. In the massed condition however, such juxtaposing effect did not exist.

The second explanation for the interleaving effect in induction is the role of attention. It is suggested that that massing impairs learning by reducing the amount of attention people pay to repeated presentations, because the massed items become highly familiar [20]. Further, according to the attention attenuation hypothesis [15, 16], recall of massed items is impaired because it is difficult to pay full attention to the second (and subsequent) presentations of massed items. In inductive learning, there is a possibility that participants might have believed that massed presentation of exemplars from a particular category made it easier for them to identify a particular psychopathological category, resulting in those massed exemplars being given less attention and receiving less processing time compared to spaced (interleaved) exemplars. It is likely also that attention is weakened across massed exemplars from the same category [16]. In contrast, when the exemplars were interleaved with exemplars from several other categories, participants may have paid more attention to and more deeply processed the exemplars.

A final possible explanation for the interleaving effect in induction is the likelihood of study-phase retrieval. When an item is presented, previous presentations of the same item may be retrieved from memory and this retrieval process enhances learning [15]. It is argued that the more difficult the retrieval, the more learning is enhanced [21, 22]. In inductive learning, the exemplars from the same category are different from each other, thus likely are more difficult to be learnt. Interleaving of exemplars from several categories might have increased the difficulty of retrieval of exemplars from the same category which then enhanced induction.

The present study also revealed that performance was significantly increased over the test blocks, and this could be due to the accuracy feedback that participants received after each trial. In addition, performance was significantly better when induction was tested shortly after the study phase than after a one week's delay. A comparison of the proportion mean scores of the first test block between the two retention conditions (see Figure 1 and Figure 2) shows a decline in participants' performance when testing was delayed a week— score decreases from 0.38 to 0.23, indicating forgetting for the long retention group. Nevertheless, at test a week later, the participants' performance improved over the test blocks, showing a benefit of returning to the learning context. Memory is enhanced when the situation during study closely resembles the situation during test [23]. It is likely that context reinstatement might have helped the participants, particularly in the long-term groups, to retrieve forgotten material and further enhance their performance during the test session. In addition, participants received feedback received on each test trial, including the correct category name when they made an error. The implementation of test with feedback might have also helped the participants in the experiment.

On the post-experimental questionnaire, consistent with findings from past studies [11-17], the majority of the participants in the short-term retention condition appeared to believe that the massed presentation made it easier to recognise the nature of each psychopathological case study during the presentation or study phase although their

performance showed the opposite. It is likely that because of the consecutive presentations in the massed condition, participants developed a sense of familiarity for the exemplars of a similar nature from a particular category. This could have led participants to perceive that the massed presentation required less effort and to infer that the learning task was easier in the massed condition. The participants then may have believed that their learning outcomes in the massed condition were better than in the spaced condition and perceived massed presentation as more helpful in learning the categories. Similarly, the same judgment on massed presentation over spaced (interleaved) presentation was obtained in the long-term retention condition, which is consistent with finding from the Zulkiply and Burt's study [18]. As mentioned, Zulkiply and Burt [18] used paintings and visually presented texts as the learning materials in their experiments. The present study provided further evidence on the participants' judgment towards massing in the long-term retention condition, even when aurally presented texts were used in the presentation/study phase.

## 5. Conclusions

The exemplar interleaving effect in inductive learning has been reported for paintings and visually presented texts [18]. The present study extended the generalisabilty of the effect to aurally presented texts. With further evidence from the present study, that induction benefits from interleaved presentation over long-term retention, a systematic approach in inductive learning or category learning could be planned and implemented by educators in order to achieve the optimal benefits of the interleaving effect in inductive learning for long-term retention.

## Acknowledgement

I thank the Ministry of Higher Education (MOHE) and Universiti Malaysia Sarawak for supporting this research under Research Acculturation Grant Scheme (RAGS/SS/02(1)/901/2012(02).

## References

[1] Cepeda, N. J., Pashler, H., Vul, E., Wixted, J.T., Rohrer, D. Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin* 2006; **132**: 354-380.

[2] Donovan, J.J., & Radosevich, D.J. A meta-analytic review of the distribution of practice effect: Now you see it, now you don't. *Journal of Applied Psychology* 1999; **84**: 795-805.

[3] Melton, A.W. The situation with respect to the spacing of repetitions and memory. *Journal of Verbal Learning and Verbal Bahavior* 1970; **9**: 596-606.

[4] Dempster, F.N. Distributing and managing the conditions of encoding and practice. In: E.L. Bjork and R. A. Bjork (Eds.), *Memory: Handbook of perception and cognition,* San Diego, CA: Academic Press; 1996, p. 317-344

[5] Rea, C.P., & Modigliani, V. The spacing effect in 4- to 9- year-old children. *Memory and Cognition* 1987;**15**: 436-443.

[6] Toppino, T.C. The spacing effect in preschool children's free recall of pictures and words. *Bulletin of the Psychonomic Society* 1993; **31**: 27-30.

[7] Ausubel, D. P. Early versus delayed review in meaningful learning. *Psychology in the Schools* 1966; **3:** 195-198.

[8] Cepeda, N.J., Coburn, N., Rohrer, D., Wixted, J.T., Mozer, M.C., Pashler, H. Optimizing distributed practice: Theoretical analysis and practical implications. *Experimental Psychology* 2009; **56**: 236–246.

[9] Bahrick, H.P., Phelps, E. Retention of Spanish vocabulary over 8 years. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 1987; **13:** 344–349.

[10] Murphy, G. L. *The big book of concepts*. Cambridge, MA: MIT Press; 2002

[11] Kang, S. H. K., Pashler, H. Learning painting styles: Spacing is advantageous when it promotes discriminative contrast. *Applied Cognitive Psychology* 2012; **26:** 97-103.

[12] Zulkiply, N., Burt, J.S. The exemplar interleaving effect in inductive learning: Moderation by the difficulty of category discriminations. *Memory and Cognition* 2013; **41**: 16-27.

[13] Kornell, N., Bjork, R. A. Learning concepts and categories: Is spacing the "enemy of induction"? *Psychological Science* 2008; **19**: 585-592.

[14] Vlach, H. A., Sandhofer, C. M., Kornell, N. The spacing effect in children's memory and category induction. *Cognition* 2008; **109**: 163-167.

[15] Kornell, N., Castel, A. D., Eich, T. S., Bjork, R. A. Spacing as the friend of both memory and induction in young and older adults. *Psychology and Aging* 2010; **25**: 498-503.

[16] Wahlheim, C.N., Dunlosky, J., Jacoby, L.L. Spacing enhances the learning of natural concepts: An investigation of mechanisms, metacognition, and aging. *Memory and Cognition* 2011; **39**: 750-763.

[17] Zulkiply, N., McLean, J., Burt, J. S., Bath, D. Spacing and induction: Application to exemplars presented as auditory and visual text. *Learning and Instruction* 2012; **22:** 215-221.

[18] Zulkiply, N., Burt, J.S. Inductive Leaning: Does interleaving exemplars affect long-term retention? (A working paper)

[19] Goldstone, R. L. Learning to perceive while perceiving to learn. In: R. Kimchi, M. Behrmann, & C. Olson (Eds.), *Perceptual organization in vision: Behavioral and neural perspectives*, Mahwah, NJ: Lawrence Erlbaum Associates; 2003, p. 233–278

[20] Hintzman, D. L. Theoretical implications of the spacing effect. In R. L. Solso (Ed.), *Theories in cognitive psychology." The Loyola Symposium.* 1974; 77-99; Potomac, MD: Erlbaum.

[21] Bjork, R. A., & Allen, T. W. The spacing effect: Consolidation or differential encoding. *Journal of Verbal Learning & Verbal Behavior* 1970; **9:** 567–572.

[22] Krug, D., Davis, B., Glover, J. A. Massed versus distributed repeated reading: A case of forgetting helping recall? *Journal of Educational Psychology* 1990; **82**: 366–371.

[23] Fisher, R. P., Craik, F. I. Interaction between encoding and retrieval operations in cued recall. *Journal of Experimental Psychology: Human Learning and Memory* 1977; **3**: 701-711.

## Appendix A. Examples of Case Studies from Experiment 2

Category TEM (Schizophrenia Disorder)

Sample 1:

Wills, 35 years old, is a successful businessman but lately, his behavioural changes seemed to affect his relationship with clients. Since 6 months ago, he had begun to hear voices that tell him he is not a good man. He has begun to talk to himself about how bad he is during meetings with clients. This has affected his relationship with clients. At the office, his workers were shocked by his very rapid changing mood, from happy to sad to angry, for no apparent reason. When he talked, it seemed that he was having thought disturbances, as he mixed up unrelated issues and could not connect his thoughts logically. He also keeps rolling up his tongue and that is somewhat annoying to his workers.

Sample 2:

Melinda, a 40-year old lady complained to her neighbours that she was fearful, depressed, and couldn't get off to sleep at night. She said she had been seeing her late mother lately, and her mother told her that her husband was going to hurt her badly. Melinda's husband was confused with Melinda's unusual behaviour, such as staring at him and locking herself in another room at night to avoid him. Two weeks later, Melinda ran away and stayed with her friend. While there, she wrote a letter to her husband saying that she was protected by a superpower and can never be hurt by anybody.

Category PLIQ (Attention Deficit Disorder- hyperactive and impulsive type)

Sample 1:

Ben, 11 years old, is a cheerful child, who often has problems in concentrating and following instructions by his teacher at school. When he does his schoolwork, he will make one or two scribbles on it and then he will start to giggle and whisper with his classmates. At home, Ben often fails to complete the house work assigned to him by his parents. For exemplar, when asked to clean up his room, he does it for a minute and then does something else which will also left unfinished. Lately, Ben also complaints that he often feels hot and he drinks more water than he usually does.

Sample 2:

Maria, age 4, had problems at preschool. Her teacher said that she seemed disorganized and inattentive when performing school activities. When she drew something, her teacher had to repeat instructions, and Maria always left half-finished drawings all over her classroom. When she played at a puzzle, she did it half way, and then she left the incomplete work for other activities. Another thing that has become apparent in Maria lately is that her hands sweat a lot though she was just doing relaxing activities. Her teacher also found that she can get very angry if she can't get what she wants, e.g., when she wants to play on the swing but the swings are all occupied.